Baysun dialekt corpus: Genesis (Baysun district is in the case of "J" dialect)

Xolova Muyassar Abdulhakimovna¹

¹Teacher of the Department of Uzbek linguistics at Termez State University, Uzbekistan Email: <u>xolova_m@umail.uz</u>

ABSTRACT

The article about the national corpus and Baysun vernacular of the local population expressed opinions about specific issues units computerize.

Keywords: dialectological corps, economic, political, and economic everyday life, computerize texts, remote exchange of information.

1. INTRODUCTION

In World linguistics, the direction of dialectology has its stages of development. If the study of syllables began with the work "Devonu lexicon Turkish" of Mahmud Kashgari (XI century)in Turkology, interest in the live speech of the people was formed in Germany in the late seventeenth century as a science of dialectology in general. At the end of the XIX beginning of the XX century, the first samples of dialectological atlases of the XX were created. From that time on, local folk dialects living on the territory of Germany were collected and compiled dictionaries. "German language Atlas "(1826-1926 years)," German lexicological Atlas "(Marburg German dialectological Institute)," French language Atlas" (12 volumes, 1902-1912 years) created by Edmon and Jilyeron; published by Yud and Yaberg [4,16].

The work on the study and scientific justification of Uzbek dialects was carried out by "Young Mathematicians" in the second half of the XIX century and was considered to be a natural, non-stop, the self-developing component of the National Language[2, 4]. It's him.D.Polivanov, I.I.Zarubin, V.V.Reshetov K.K.Yudexin, A.K.Borovkov, Gazi Scientist Yunusov, F.Zire "Lost In Test MatchShaabdurahmonov, A.Shermatov, N Rajabov, X.Doniyorov, Q.Muhammadjonov, D.Abdurahmonov, B.Jo'reyev, A.Jo'rayev, A.Isayev, N.Shoimova, S.Rahimov, X.Fayziyeva, N.The services of such scientists as Murodova in the development of Uzbek dialectology are incomparable.

In our linguistics within the framework of studies on phonetic, lexical, grammatic levels of Uzbek dialects, in the next 50 years, "Khorezm dictionary dialects" (1961). "Lexicon of Uzbek dialects" (1966). "Lexicon of dialects of Tashkent region" (1966). "Lexicon of Kokand group dialects" (1971). "Oguz dialect of Uzbek language" (1973). "Dictionary of Uzbek folk dialects" (1971), "lexicon of Uzbek dialects in Surkhondarya region " (1974), "Uzbek dialects in Karakalpak'istan" (1977), lexicon of Kipchak dialects (1979), "Uzbek dialects of Surkhondarya region" (1985), Zhurayev B. Studies such as" higher Kashkadarya Uzbek dialects " have been carried out.

2. MAIN BODY

Attention to Boysun dialectal units, Zarubin divided Uzbek dialects into four groups, not taking into account Kipchak dialects living in a large part of the territory of Uzbekistan, Yudakhin first classifies Uzbek dialects according to their relationship with other (Tajik) dialects, then, having met mufassal, the dialects into 5 groups: 1)Tashkent, 2) Farg'ona, 3) Kipchak, 4) Khiva, (Khiva), 5) northern Uzbek dialects the accent is included in the Gypsy group. Prof. E. D. Polivanov in his classification turns out to be organized and non-organized groups, taking into account the monetization (of fraternal languages) and hybridization (structurally, of different languages). The dialect of the regions under the influence of near the centre of the Boysun district is included in the dialect of Samarkand-Bukhara type 1, and the dialect of the third dialect of the dialect "ZH" is included in the middle Uzbek dialect.

Prof. Sh. Shoroabdurahmonov points out that the theory of Polivanov is explained to the outside world, does not take into account the results of internal development that occurred in the dialects of the Uzbek language, and there is confusion in some theoretical conclusions. According to the Borovkov classification of the Boysun dialect-Uzbek or "j" type of dialects [4,35], Uzbek-Kipchak group according to the classification of geese scientists, V. V. According to the Reshetov classification, it belongs to the Kipchak dialect of the north-west group of the Uzbek language [4,36]. In the process of language development, it's content (lexical) has undergone significant, serious changes compared to other areas [9, I, 3-b]. As part of the Explanatory Dictionary of the Uzbek language, more than 80 000 lexical interpretations of the national language and dialectal stable units, which demonstrate the identity of the people, are widely covered by the form and content of linguistic units in our language.

The events that occur in the life of any people are initially reflected in the lexicon of the language. Therefore, the lexicon is the most mobile in the elements of the language structure [6, 23] and the window of the people's household level [1,72]. The Uzbek people experienced very difficultly and different stages in the process of historical development were in social, political and cultural relations with different peoples and enriched their vocabulary. Therefore, the composition of the dictionary of the Uzbek language is the product of a long period of time, it has very rich vocabulary and word forms, phraseological units, which are understandable for all members of the same people, are used a lot in economic and social life, are able to give a clear meaning, are the basis for word-making and-making.

It is appropriate to study the syllables again after a certain period. Although many books, monographs and articles on phonetics, lexicons, semantics of Uzbek dialectal vocabulary should be kept in the focus of attention of researchers, research on dialectal vocabulary should still be carried out in the present day, considering that the development of technology and the growing cultural and household life in this period of rapid development of the language of the future generation .

Today it is difficult to imagine the development of all spheres of socio-economic life without computer technology, and at the same time in the solution of the issues facing linguistics, too, gradually began to resort to electronic text corpses, which were formed with the help of computer technology. The body of electronic texts, it consists of a certain collection of texts, putting them into a Specially defined norm, structure and comparing all alternatives of different socio-territorial, functional and year-cross-section of the language from Linguistic Science Information we find, the role of this system for studies is incomparable [8,4-b].

Research on the first linguistic Corps in science began to appear in the 60-of the last century. In 1963, at the University of Braun, USA, the first large mantle Corps was created in the storage car (Brown Corps). Text authors U. Francis and H. The American variant of 500 prose texts, consisting of 50 genres, 2000 words in English, was developed by the dialectism [8,4-b]. In World linguistics, the creation of language bodies is carried out in Austria, China, Germany, Portugal, Finland, Scandinavia, Poland, Lithuania, Georgia (for example COR-DIAL-SIN; Helsinki corpus of English dialects; The Nordic Dialect Corps; Archive Deutsch, Die Dialektadebank; LEXDIALGRAM etc.), Specific dialectal corpses (Dialectium podcorpus v sostave Nationalnogo corpusa russkogo yazika, Saratovsky dialectologicheskiy Corpus, Elektronnaya biblioteka russkix narodnix govorov, Elektronniy Corpus dialectnoy kulturi Kubani I dr.), specific dialectal Corpus (dialect podcorpus v sostave Nationalnogo Corpus russkogo yazika, Saratovsky dialectologicheskiy Corpus, Elektronnaya biblioteka russkix narodnix govorov, Elektronniy Corpus dialectnoy kulturi i dr.) [5]it is worth noting the research conducted by scientists of such countries.

In Russian linguistics, a special school was created in the creation of the national body: the structure of the Russian national body («Национальный корпус русского языка»), Sankt-Peterburg, Saratov the work done in the institutes and schools of the Tomsk dialectological Corps is proof of our opinion[7].In Uzbek linguistics, computer linguistics, lexicographic processing of text, Linguostatic analysis of the text. Polotov, M.Ayimbetov, S. Muhammedova, S.Karimov, A. Babanarov, D.Surinboyeva, N.Abdurahmonova, A.Research was conducted by such scientists as Norov, a philologist by Hamroyeva on the topic "linguistic foundations of the creation of the Uzbek language author's body". science. Ph. D. in Philosophy (PhD) in business, a. Eshmo'minov's "database of synonyms of the national corpus of the Uzbek language" (philologist. science. one of the practical steps in computer linguistics is the research carried out on the topic of philosophy doctor (PhD)) [3].

The body of language or texts is a framework that perfectly reflects the solution of issues facing a computer format, structural, cosmetic, linguistics. A special feature of the modern corpus is that the metallurgical information contained in it has a specified format and carries out the presence of a text selection. The set of norms, the sum of search engine texts with a sign of principles, is a linguistic Corps[8,5 - b]. Sometimes the sum of the desired texts [subjects (family, history, culture, etc.)], which acquires a common unity.), genres (bibliographic, event study, interview, etc.), conversation (dialogic,),

information (accurate, estimated, subjective, etc.).)] the term Corps is given. Analysis of objectives from the creation of Corpus texts is given as follows:

- The composition of the text is the application of linguistic Language units in the sentence;
- Sufficient competence of units in a large linguistic Corps;
- One-time or multiple-time participation of the created language Corps in the solution of problems facing linguistics. For example, in the processes of lexical-grammatical analysis or the problems of the structure of the components of lexical dictionaries.

I.D.Chesnokova va M.EManshinlar "Rus milliy korpusi lingvistik tadqiqotlarning vositasi" his article highlights the research on the different sections of the language Corps: 1. Main body. Two. Corps of deep syntactic analysis. The three. Gazeta Corps. The four. Parallel text Corps. The five. Dialectical text Corps (texts written by dialectologists from the ethnic population residing in the area). The six. Corps of poetic texts. The seven. Language teaching corps. Eight. Verbal colloquial Corps. The 9. Axentological (female) Corpus. Ten. Multimedia housing. The 11. Historical language Corps[10].

If we approach the issue of creating a dialectal body, then the materials collected during the expeditions in themselves are a reflection of modern forms of the dialect of the national language, unequal sources of historical and classical literature and the conversational culture of the ethnic population. Boysun's dialect units on the basis of dictionaries can be of assistance not only for solving important problems facing theoretical and methodical linguistics researchers but also for research in various aspect. These are the problems of traditional and oral colloquial speech categories of speech culture, the inclusion of language units in the dictionaries of the Uzbek language, lexical-semantic (motivational, emotional-affective, evaluative, levelling, intensity) categories together with studies that determine the specific outlook of the owner of the spoken language in the dialect, the scope of application of language units.

So far, even though scientists in the territory of Uzbekistan have created a Dialictism through the results of the expedition, the problems of creating an electronic form that reflects the texts of the dialect and not only Boysun, but also other dialects in general in the territory of our country have not been solved, show the relevance of our work. Sheva is a means of ensuring communication of people belonging to a certain region, and the scope of its application is limited. On the one hand, the creation of the Uzbek language Corps provides an easy solution to the issues related to the study, comparison, application of dialectal materials and finding their alternatives in other languages, on the other hand, proving that the existing units in the language are not lost and reflect the national sentiment in it.

In recent years, the results of dialectal tours, organized by students, remain in the records. Modern scientific research shows that we need to create a comprehensive, easy-to-use and easy-to-use electronic source, dialectal body. Currently, the establishment of the dialectal Corps is considered one of the urgent tasks facing the Uzbek linguistics, the implementation of the work aims to electronic the materials collected during the expedition (words reflecting economic-political, economic-household life, fairy tales, Proverbs, metals) and the dictionaries compiled through them, texts reflecting the dialect units.

The first step in carrying out the study should be the following:

- 1. Through the introduction of units, words, sentences, texts into the computer in the form of handwritten syllables, create electronic texts and files.
- 2. To perform this work, set the value of all the handles and save them in the form of a picture and a DJVU file.
- 3. The goal is to develop a single symbolic character, which should be phonetic, morphological and lexical-grammatical analysis of texts on the dialect of Boysun, having set it before us.
- 4. To give national expressions and vocabulary, which have in common with artistic literature, subject to dictation norms.
- 5. Analysis of phonetic and morphological characteristics of nationwide, dialectal, historical-dialectal words and vocabulary.
- 6. Indicate the origin, development, alternation of dialectal characteristics of words specific to dialect.
- 7. Evaluation of the relationship of dialects with other languages.
- 8. To determine the participation of other languages in the composition of syllables.

3. CONCLUSION

In conclusion, it aims to find answers to a wide range of language issues, regardless of the problems and tasks put before linguistic research. The use of texts, genres written in different styles, as well as questionnaires conducted through questionnaires will ensure a reliable course of research. Along

with the researcher's scientific attitude to the collected materials, he also collected them on the cards and made up the brought them into the electronic form on the computer. Technical means are used when storing voice-collected data. Taking into account the fact that the scientists collected data collected by the scientist are left in placing them in the computer memory, the use of stored comparative and common language units, words, vocabulary and text corpus reduces the waste of time for research, increasing the level of knowledge of research within the framework of linguistics. Now that young researchers can interact with the entire database via the internet, remote relationships are a great opportunity for them. All this will serve as a new stage for research in Uzbek linguistics.

REFERENCES

- 1. Abayev V.I. O razvitii literaturnix yazikov iranskix narodov Sovetskogo Soyuza. V. sb.: Zakonamernosti razvitiya literaturnix yazikov SSSR v sovetkuyu epoxy. Moskva,1969.
- 2. Ashirboyev S. O'zbek dialektologiyasi. -Toshkent, 2004.
- 3. Eshmo'minov A. "Ozbek tili milliy korpusining sinonim so'zlar bazasi" filolog. fan. bo'yicha falsafa doktori (PhD)...diss. Avtoref. –Qarshi. 2019.
- 4. Reshetov V.V., Soabdurahmonov Sh. O'zbek dialektologiyasi. –Toshkent, 1978.
- 5. Y.B.Ivansova. Tomskiy dialektniy korpus: obosnovaniye kosepsii I perspektivi razvitiya. "Voprosi leksikografii-Russian Journal of Leksicograpfhy", 2017, 11, 54-70-str
- 6. Vinogradov B. B. Razvitiya sovetskogo yazikoznaniya. Izd. APN. -Moskva, 1950.
- 7. Yurina Y. A. Tomskiy dialektniy korpus: v nachale puti. Vestnik Tomskogo universiteta №2 (14) 2011.
- 8. Захаров В.П. Корпусная лингвистика: Учебно-метод. пособие. –СПб., 2005. 48 с.
- 9. O'zbek tilining izohli lug'ati I-V tomlar, -Toshkent, 2006-2008.
- Chesnokova. I.D. Manshin M.Y. Natsionalniy korpus russkogo yazika kak osnovnoy instrument poiska lingvisticheskix issledovaniyax (na primere poiska antonimov v publitsisticheskix tekstax). Izvestiya VGPU filologicheskiye nauki, 2018.